

ND

OM of: US-09-784-340-2 to: EST: * out_format : pfs
Date: Aug 27, 2001 7:41 PM

About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:

-MODEL=frame+_p2n.model -DEV=xlp
-O=/cgn2.1/USPTO_spool/US09784340/runat_27082001.123146_78/app-query.fasta.1.591
-DB=EST -QFMT=fastap -SUFFIX=olip2n.rst -GAPOP=4.500
-GAPEXT=0.050 -MINMATCH=0.100 -LOOPEXT=0.000 -LOOPEXT=0.000
-GAPOP=4.500 -GAPEXT=0.050 -XGAPOP=60.000 -XGAPEXT=60.000
-DELOP=6.000 -XGAPEXT=7.000 -XGAPOP=60.000 -XGAPEXT=60.000
-DELDP=6.000 -DELXEXT=7.000 -START=1 -MATRIX=oligo
-TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR_SCORE=quality
-THR_MIN=1 -ALIGN=1 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -MINLEN=0
-MAXLEN=2000000000 -USER=US09784340.@cgn1.1.3561 -NCPU=6
-ICPU=3 -LONGLOG -NO_XLPHY -WAIT -THREADS=1

Search information block:

Query: US-09-784-340-2

Query length: 527

Database: EST: *

Database sequences: 10728115

Search time (sec): 1332.720000

WARN: XGAPOP and YGAPOP must be equal. Assuming YGAPOP=XGAPOP=60.000
WARN: XGAPEXT and YGAPEXT must be equal. Assuming YGAPEXT=XGAPEXT=60.000

score_list:

Sequence	Strd Orig	ZScore	EScore	Len	Documentation
gb_est198:BG249053	83.00	1302.32	1.5e-74	671	! BG249053 602361828f1 NIH_MGC_89
gb_gss8:AO206472	74.00	1338.91	1.9e-65	501	! AO206472 HS_324LA2_G09_T7 CT
gb_gss36:BB0712	54.00	971.77	5.2e-45	444	! BB0712 CT-HSP-2050K23.TF CT-H
gb_gss10:AO311717	44.00	785.45	1.2e-34	601	! AO311717 RPI11-155015.TV RPI1
gb_est18:AI303265	41.00	731.26	1.2e-31	525	! AI303265 u175908.x1 Sugano mons
gb_est25:AI791045	41.00	731.26	1.2e-31	525	! AI791045 uk55a11.y1 Sugano mons
gb_est24:AI788285	41.00	730.47	1.4e-31	583	! AI788285 uk56b10.y1 Sugano mons
gb_est16:AI158921	41.00	729.46	1.6e-31	666	! AI158921 uF05C02.x1 Sugano mons
gb_est21:AI529553	41.00	727.74	2.0e-31	836	! AI529553 u182a02.x1 Sugano mons
gb_hic:AK008601	41.00	720.76	5.0e-31	2099	! AK008601 Mus musculus adult ma
gb_est100:BG402144	37.00	653.58	2.8e-27	898	! BG402144 60245689f1 NIH_MGC_75
gb_est83:BF080415	33.00	585.22	1.8e-23	449	! BF080415 231292 MARC_2P1G Sus s
gb_est14:AF202345	33.00	584.67	1.9e-23	483	! AF202345 AF202345 Homo sapiens
gb_est83:BF080461	33.00	583.59	2.2e-23	557	! BF080461 231360 MARC_2P1G Sus s
gb_est109:TF67857	28.00	494.03	2.1e-18	403	! TF67857 yc27h03.y1 Stratagene
gb_est40:AV696157	28.00	493.90	2.2e-18	410	! AV696157 AV696157 GRC Homo sap
gb_est40:AV655524	28.00	493.84	2.2e-18	413	! AV655524 AV655524 GRC Homo sap
gb_est14:AA962194	28.00	493.50	2.3e-18	432	! AA962194 oq69h06.s1 NCI_CGAP_K
gb_est14:AA967868	28.00	492.67	2.3e-18	482	! AA967868 oq81f04.s1 NCI_CGAP_K
gb_gss10:AO311499	28.00	492.18	2.7e-18	514	! AO311499 RPI11-12108.TJ RPI1
gb_est20:AI480418	28.00	491.95	2.8e-18	730	! AI480418 tms2b06.x1 NCI_CGAP_K
gb_est91:BF688309	28.00	489.36	3.9e-18	746	! BF688309 602185172f1 NIH_MGC_43
gb_est40:AV681832	28.00	489.16	4.0e-18	766	! AV681832 GCB Homo sapi
gb_est25:AI821970	25.00	453.38	3.9e-15	669	! AI821970 cb22c12.x5 NCI_CGAP_K
gb_est20:AI433838	25.00	437.00	3.2e-15	512	! AI433838 f117b01.x1 NCI_CGAP_K
gb_est83:BF132129	25.00	436.63	3.3e-15	538	! BF132129 601821157f1 NIH_MGC_58
gb_est13:AA891683	24.00	418.87	3.3e-14	494	! AA891683 EST195486 Normalized
gb_est13:AA893486	24.00	418.84	3.3e-14	496	! AA893486 EST197289 Normalized
gb_est14:AA945116	24.00	417.82	3.7e-14	568	! AA945116 EST200615 Normalized
gb_est53:AW916387	24.00	417.29	4.0e-14	609	! AW916387 EST147678 Rat gene ind
gb_est100:BG427877	24.00	414.04	6.1e-14	935	! BG427877 602505156f1 NIH_MGC_75
gb_est40:AV693253	23.00	403.64	2.3e-13	325	! AV693253 GAG33253 CGP Homo sap
gb_est44:AM237578	23.00	403.28	2.4e-13	341	! AM237578 km73h06.x1 NCI_CGAP_K
gb_est109:TF50951	23.00	402.38	2.7e-13	384	! TF50951 yb88904.r1 Stratagene
gb_est3:AA209027	23.00	400.15	2.8e-13	395	! AA209027 mw73d12.r1 Soares mons
gb_est61:BB283558	23.00	400.15	3.0e-13	515	! BB283558 BB283558 Riken full-1e
gb_est22:AI162758	23.00	399.98	3.7e-13	527	! AI162758 ty81c10.x1 NCI_CGAP_K
gb_est20:AI451845	23.00	399.78	3.8e-13	541	! AI451845 mx17h01.y1 Soares mons
gb_est69:BE013295	23.00	399.60	3.9e-13	554	! BE013295 123199 MARC_1P1G Sus s

seq_name: gb_est198:BG249053

seq_documentation_block: 671 bp mRNA 13-FEB-2001

LOCUS BG249053 671 bp mRNA 13-FEB-2001

DEFINITION 602361828f1 NIH_MGC_89 Homo sapiens cDNA clone IMAGE:4470199 5'

ACCESSION BG249053

VERSION BG249053.1 GI:12758869

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE 1 (bases 1 to 671)

AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgapds@mail.nih.gov

Tissue Procurement: ATCC

CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:

<http://image.llnl.gov>

Plate: LLM10287

High quality sequence stop: 661.

FEATURES

Location/Qualifiers

1..671

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:4470199"

/clone_lib="NIH_MGC_89"

/tissue_type="hypertrophied, cell line"

/lab_host="DH10b (phage-resistant)"

/note="Organ: Kidney; Vector: pCMV-Sport6; Site: 1: Not; Site: 2: Salt; Cloned unidirectionally; oligo-dT primed. Average insert size 1.3 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC library."

BASE COUNT 206 a 123 c 153 g 189 t

ORIGIN

alignment_scores:

Quality: 83.00 Length: 83

Ratio: 1.000 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-784-340-2 x BG249053

Align seg 1/1 to: BG249053 from: 1 to: 671

91 ValAspleuAlalaAsnValleuProglyLeuSerThrTrpInserva 107

|||||

111 GPTACCTGACCTGGAAGTCTCCACAGCTTATCAACCTGCATACAGT 160

|||||

107 IILeLysLeuSnsPhePheValGluIleArgGlyThrLeuLysMetM 124

|||||

161 TATTAATTAATGATTTTGTGTAATTAAGACGAACTTAATAATGA 210

|||||

124 eTcysGluSerPheIleTyraSngInThrLeuMetLysLysLeuGln 140

|||||

211 TGTGTGAGAGCTTATATCAATACAGAGCTTATGAGAGAACTACAGGA 260

```
141 ThrAsnTyArpValMetLeuIleAspProValIleProCysGlyAspLe 157
|||||
261 ACCAACTACATGTATGCTTATAGACCCGTGATTCCTCGTGAACCT 310
|||||
157 uNcctAlaGluLeuLeuAlaValProPheValLeuThrLeuArgIleSer 173
|||||
311 GATGCTGAGTGTCTGCAGTCCCTTTGTGTGCACACTTAGAATTTC 359
|||||

seq_name: gb_gss8:AQ206472

seq_documentation_block:
LOCUS      AQ206472      501 bp      DNA      GSS      17-SEP-1998
DEFINITION HS_3241_A2_G09_T7 CIT Approved Human Genomic Sperm Library D Homo
sapiens genomic clone Plate=3241 Col=18 Row=M, DNA sequence.
ACCESSION  AQ206472
VERSION     AQ206472.1  GI:3617042
KEYWORDS    GSS.
SOURCE      human.
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE   1 (bases 1 to 501)
AUTHORS     Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
Hood,L.
TITLE       Sequence-tagged connectors: A sequence approach to mapping and
scanning the human genome
JOURNAL     Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)
MEDLINE     99380589
COMMENT     Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Sequence Tagged Connector
Plate: 3241 Row: M Column: 18
Class: BAC ends
High quality sequence stop: 501.
Location/Qualifiers
1..501
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="CIT Approved Human Genomic Sperm Library D"
/sex="male"
/note="Organ: sperm; Vector: pBelobAC11; BAC Clones in
E-Coli DH10B"

BASE COUNT  135 a 110 c 93 g 160 t 3 others
ORIGIN
```

```
FEATURES
source
1..501
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="CIT Approved Human Genomic Sperm Library D"
/sex="male"
/note="Organ: sperm; Vector: pBelobAC11; BAC Clones in
E-Coli DH10B"

BASE COUNT  135 a 110 c 93 g 160 t 3 others
ORIGIN
```

```
alignment_scores:
Quality: 74.00 Length: 74
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
```

```
alignment_block:
US-09-784-340-2 x AQ206472 ..
```

```
Align seg 1/1 to: AQ206472 from: 1 to: 501
```

```
454 ProLeuAspArgAlaValPheTrpIleGluPheValMetArgHisIysG1 470
|||||
46 CCCCTAGATCGACAGCTTCTGATCGAGTTGTGTCATCGCCACAAGG 95
|||||
470 yAlaLysHisLeuArgSerAlaIleHisAspLeuThrTrpPheGlnHisT 487
|||||
96 AGCGAAGCACCTCGATCGACGCTGCCATGACCTCCACCTGTTCCACACT 145
|||||
487 ySerIleAspValIleGlyPheLeuLeuThrcysValAlaThrAlaIle 503
|||||
```

```
|||||
146 ACTCTATAGATGATGATTGGTTCCTGACCTGATGTCGCACTGCTATA 195
|||||
504 PheLeuPheThrLysCysPheLeuPheSerCysGlnLysPheAsnLysTh 520
|||||
196 TCTCTGTTCAAAAGATGTTTATTTTCTCTCAAAAATTAATAAAGAAC 245
|||||
520 rArgLysIleGluLysArgGlu 527
|||||
246 TAGAAAGATAGAAAGAGCGAA 267
|||||
```

```
seq_name: gb_gss36:B80712

seq_documentation_block:
LOCUS      B80712      444 bp      DNA      GSS      24-OCT-1998
DEFINITION CIT-HSP-2050M23.TF CIT-HSP Homo sapiens genomic clone 2050M23, DNA
sequence.
ACCESSION  B80712
VERSION     B80712.1  GI:2867735
KEYWORDS    GSS.
SOURCE      human.
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE   1 (bases 1 to 444)
AUTHORS     Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Baas,S., Linher,K.,
Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
Simon,M. and Venter,J.C.
TITLE       Use of a random BAC End Sequence Database for Sequence-Ready Map
Building (1998)
JOURNAL     Unpublished (1998)
COMMENT     other_GSSs: CIT-HSP-2050M23.TR
Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdamas@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tdb/hungen/bac_end_search/bac_end_search.html
Seq primer: MJ3-21
Class: BAC ends.
```

```
FEATURES
source
1..444
/organism="Homo sapiens"
/db_xref="GDB:7056698"
/db_xref="taxon:9606"
/clone_lib="2050M23"
/clone_lib="CIT-HSP"
/sex="male"
/cell_type="Sperm"
/note="Vector: pBelobAC11; Site_1: HindIII; Site_2:
HindIII"

BASE COUNT  124 a 110 c 71 g 139 t
ORIGIN
```

```
alignment_scores:
Quality: 54.00 Length: 54
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
```

```
alignment_block:
US-09-784-340-2 x B80712 ..
```

```
Align seg 1/1 to: B80712 from: 1 to: 444
```

```
434 SerSerTyLysGluAsnAlaMetArgLeuSerArgIleHisAspG1 450
|||||
282 TCCAGTTATTAAGAGATATGATGATTATATCAAGATTCACCATGATCA 331
|||||
```



```

/sex="female"
/dev.stage="adult"
/lab_host="DH10B"
/note-Organ: Kidney; Vector: pME18S-FL3; Site_1: DraIII
(CACTGTG); Site_2: DraIII (CACCAGTGT); 1st strand cDNA
was primed with an oligo(dT) primer
[ATGGCGCCCTTTTCTTTTTTTTTTTTTTTT]; double-stranded cDNA was
ligated to a DraIII adaptor [TCGTGGCCTACGG], digested
and cloned into distinct DraIII sites of the pME18S-FL3
vector (5' site CACTGTG, 3' site CACCAGTGT). XhoI should
be used to isolate the cDNA insert. Size selection was
performed to exclude fragments <1.5kb. Library
constructed by Dr. Sumio Sugano (University of Tokyo
Institute of Medical Science). Custom primers for
sequencing: 5' end primer CTCTGCCTTAAGAAGCTGGC and 3' end
primer CGACCTGCAGCTCGAGCA."
BASE COUNT      146 a      130 c      154 g      152 t      1 others
ORIGIN

alignment_scores:
    Quality:      41.00          Length:      41
    Ratio:        1.000         Gaps:       0
Percent Similarity: 100.000     Percent Identity: 100.000

Alignment_block:
US-09-784-340-2 x AT788285 ..

Align seg 1/1 to: AT788285 from: 1 to: 583

435 SetTyTLyGLuaSnAlaMetArgLeuSerArgILleHisHisAspGlnPr 451
|||||
177 TCCTAAAGAGAGAAATGCCATGCGGTATCAAGAAATTACCATGATATCGCC 226
|||||
451 oVallYSProlEuaSPARGalavalApherPIlegluPheValMetArgh 468
|||||
227 AGTGAAGCCCCCTTGACCGACGACACTCTTCTGTGATTGAGTTGTATGCGTC 276
|||||
468 IsLySGlYalalYSHisLauArg 475
|||||
277 ACMAAGAGCCCAAGCATCTTGGT 299
|||||

seq_name: gb_estt16:AII58921

seq_documentation_block:
LOCUS      AII58921                666 bp            mRNA           EST           02-OCT-1998
DEFINITION uF05c02.x1 Sugano mouse liver m1a Mus musculus cDNA clone
IMAGE:1499714.3' similar to gb:U06641 UDP-GLUCURONOSYLTRANSFERASE
288 PRECURSOR, MICROSMAL (HUMAN); gb:X06358 Mouse mRNA for
UDP-glucuronosyltransferase (MOUSE); mRNA sequence.
ACCESSION  AII58921
VERSION    AII58921
KEYWORDS   EST.
SOURCE     house mouse.
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE  1 (bases 1 to 666)
AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,R., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Watsonston,R.
            The Mashu-HMI Mouse EST Project
            Unpublished (1996)
COMMENT    Contact: Marra M/Mouse EST Project
            Washu-HMI Mouse EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LML ; contact the

```

IMAGE Consortium (infoimage.llnl.gov) for further information.
MGI:997338
Seq primer: custom primer used
High quality sequence stop: 464.
location/Qualifiers
1. 666
/organism="Mus musculus"
/strain="C57BL"
/db_xref="taxon:10090"
/clone="IMAGE:1499714"
/clone_id="Sugano mouse liver: mlia"
/sex="female"
/dev_stage="adult"
/lab_host="DH10B"
/note="Organ: liver; Vector: pME18S-FL3; Site:1: DraIII
(CACCTGTG); Site:2: DraIII (CACCATGTG); 1st strand cDNA
was primed with an oligo(dT) primer
[ATGCGCCCTTTTCTTTTCTTTTCTTTT]; double-stranded cDNA was
ligated into distinct DraIII sites of the pME18S-FL3
vector (5' site CACCTGTG, 3' site CACCATGTG). XhoI should
be used to isolate the cDNA insert. Size selection was
performed to exclude fragments <1.5kb. Library
constructed by Dr. Sumio Sugano (University of Tokyo
Institute of Medical Science). Custom primers for
sequencing: 5' end primer CTCTGCTCTTAAAGCTGCG and 3' end
primer CGACCTGCGACCTGCGACAC."

BASE COUNT 178 a 144 c 139 g 203 t 2 others
ORIGIN

alignment_scores:
Quality: 41.00 Length: 41
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-784-340-2 x AII58921/rev ..

Align seg 1/1 to reverse of: AII58921 from: 1 to: 666

seq_name: gb_est21:A1529553

seq_documentation_block:
LOCUS A1529553 836 bp mRNA EST 18-MAR-1999
DEFINITION u182a02.x1 Sugano mouse liver mlia Mus musculus cDNA clone
IMAGE:1888874.3' similar to gb:U06641 UDP-GLUCURONOSYLTRANSFERASE
288 PRECOURSOR, MICROSMAL (HUMAN); gb:U06358 mouse mRNA for
UDP-glucuronosyltransferase (MOUSE); mRNA sequence.
ACCESSION A1529553
VERSION A1529553.1 GI:4443688
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 836)
Marrin, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T.,
Underwood, K., Stepien, M., Theising, B., Allen, M., Bowers, Y., Peterson
, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurr, R., Riltter
, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R.,

ELGARGHEVTLVKYPSIIIDSKRIPIPHENIPLEYEIEFENRNLNLANVNIPI
 ISLMAKATLOPEFLQVTCDESDICSVYNOKEMKILDAQYDVVVIDPVDCGSL
 AVLQIPVYTLRFSGMYMEKHGOLPIPLTIVYVMSLTDNMFTFRVKNMFSL
 LFEWLMQYDFAMQYFSETLGRPTTCKTVEADWILRTYWDVEFPVPLPNEF
 VGLGCKPAKPLPKEMEERVOSSGEHVVVSLGSVMKNTLTKANLISVLAQIDOK
 VMRRSGKKPALIGSNTLRFNFNIPONDILGHRTKAFITRGTNGIYEALHYGVPMG
 VPMIDOPNINAHMEAKKALKVSISTMTSTDLISVRVAINPSPYKEMAMRLSRHH
 DQPVKPLBRAVRIEIEVVRHKKAKHLRVAAHDLISRFQYHSLDVIGLILCVVTLTFIT
 TRFCELVCKLVMKESKMKNNRKKNN"

BASE COUNT 618 a 406 c 467 g 608 t

ORIGIN

alignment_scores:
 Quality: 41.00 Length: 41
 Ratio: 1.000 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
 US-09-784-340-2 x AK008601 ..

Align seg 1/1 to: AK008601 from: 1 to: 2099

435 SerTyrIySGluAsnAlaMetArgLeuSerArgIleHisAspGlnPr 451
 ||||||||||||||||||||||||||||||||||||||||||||||||
 1323 TCTTATTAAGAGATGCCATCGGTTATCAAGAAATCCACATGATCAGCC 1372
 451 oValIySProLeuAspArgAlaValPheTrpIleGluPheValMetArgH 468
 ||||||||||||||||||||||||||||||||||||||||||||||||
 1373 AGTGAACCCCTGGACAGACAGCGTCTTGATGATGATTCATCGCTC 1422
 468 IAsIySGlyAlaIySHisLeuArg 475
 ||||||||||||||||||||||||||||||||||||||||||||||||
 1423 ACAAGAGACCAACACATCTTCT 1445

seq_name: gb_est100:BG402144

seq_documentation_block:

LOCUS BG402144 898 bp mRNA EST 12-MAR-2001
 DEFINITION 602465689F1 NIH_MGC_75 Homo sapiens cDNA clone IMAGE:4593694 5',
 mRNA sequence.

ACCESSION BG402144
 VERSION BG402144.1 GI:13295592
 KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: CLONTECH Laboratories, Inc.
 cDNA Library Preparation: CLONTECH Laboratories, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 plate: LICM134 row: a column: 23
 High quality sequence stop: 557.

FEATURES
 SOURCE location/Qualifiers
 1..898
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4593694"
 /clone_lib="NIH_MGC_75"
 /lab_host="DH10B (r1 phage-resistant)"
 /note="Organ: Kidney; Vector: pDNK-LIB (Clontech); Site: 1:
 SfiI (ggcgagctcgcc); Site 2: SfiI (ggcgatcagcc); 5' and
 3' adaptors were used in cloning as follows: 5' adaptor

sequence: 5'-CACGGCATTATGCCC-3' and 3' adaptor sequence:
 5'-ATTCTAGAGCGCGAGCGCGCGACAG-dt(30)BN-3' (where B = A,
 C, or G and N = A, C, G, or T). Average insert size 1.65
 kb (range 0.5-4.0 kb). 15/15 colonies contained inserts
 by PCR. This library was enriched for full-length clones
 and was constructed by Clontech Laboratories (Palo Alto,
 CA). Note: this is a NIH_MGC Library."

BASE COUNT 249 a 175 c 165 g 309 t

ORIGIN

alignment_scores:
 Quality: 37.00 Length: 37
 Ratio: 1.000 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
 US-09-784-340-2 x BG402144 ..

Align seg 1/1 to: BG402144 from: 1 to: 898

491 ValIleGlyPheLeuLeuThrCysValAlaThrAlaIlePheLeuPheTh 507
 ||||||||||||||||||||||||||||||||||||||||||||||||
 1 GTGATTGGGTTCTCTGCTGACCTGTGTGCGCACTGCTATATCTTTCTCAC 50
 507 rLyScySPheLeuPheSerCysGlnIySPheAsnIySThrArgIySileG 524
 ||||||||||||||||||||||||||||||||||||||||||||||||
 51 AAAAGTTTTTTTATTTTCTCTGTCGCAAAAATTTATTAACCTGAAGATVAG 100
 524 IuLySArgGlu 527
 ||||||||||||||||||||||||||||||||||||||||||||||||
 101 AAAAGAGCGAA 111

seq_name: gb_est83:BF080415

seq_documentation_block:

LOCUS BF080415 449 bp mRNA EST 18-OCT-2000
 DEFINITION 231292 MARC 2Pig Sus scrofa cDNA 5', mRNA sequence.

ACCESSION BF080415
 VERSION BF080415.1 GI:10874245
 KEYWORDS EST.

SOURCE pig.

ORGANISM Sus scrofa

REFERENCE
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.
 1 (bases 1 to 449)
 Fahnenkrug, S.C., Freking, B.A., Rohrer, J.A., Smith, T.P.L., Casas, E.,
 Stone, R.T., Heaton, M.P., Grosse, W.M., Bennett, G.A., Laegreid, W.W.
 and Keeler, J.W.

JOURNAL Design and use of two pooled tissue normalized cDNA libraries for

EST discovery in swine
 Unpublished (2000)
 Contact: Smith TPL
 USDA, ARS, US Meat Animal Research Center
 PO Box 166, Clay Center, NE 68933-0166, USA
 Tel: 402 762 4366
 Fax: 402 762 4390
 Email: smith@email.marc.usda.gov
 Single pass sequencing. Bases called and alt. trimmed with phred
 v0.980904.e. Vector identified by cross_match with the -minscore 18
 and -mismatch 12 options.

PCR primers
 FORWARD: AGGAAACAGCTATGACCAT
 BACKWARD: GTTTCGCCAGTCACGACG
 plate: 52 row: c column: 11
 Seq primer: ATTAGGTGACACTATAG.

FEATURES
 SOURCE location/Qualifiers
 1..449
 /organism="Sus scrofa"
 /db_xref="taxon:9823"
 /clone_lib="MARC 2Pig"
 /tissue_type="pooled"
 /lab_host="DH10B"

/note="Vector: PCMV SPORT6; Site.1: XbaI; Site.2: XhoI;
Library made from pooled tissue from testis, ovary,"
endometrium, hypothalamus, pituitary, and placenta."
BASE COUNT 105 a 116 c 108 g 120 t
ORIGIN

alignment_scores:
Quality: 33.00 Length: 33
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-784-340-2 x BF080415 ..

Align seg 1/1 to: BF080415 from: 1 to: 449

443 LeuSerArgIleHisAspGlnProValLysProLeuAspArgAlaVala 459
|||||
249 TTATCAAGAAATTCACACGATCAACCTGTGAAGCCCTGGACGGCGGT 298
459 lPhetPrlleGluPheValMetArgHisLysGlyAlaLysHisLeuArg 475
|||||
299 CTCTGGATTGAGTTTGTGATGCGCCACAAAGAGCCACGACCTGCGG 347

seq_name: gb_est14:AF202345

seq_documentation_block:

LOCUS AF202345 483 bp mRNA EST 01-NOV-2000
DEFINITION AF202345 Homo sapiens HepG2 Homo sapiens cDNA clone 128-7, mRNA
sequence.

ACCESSION AF202345
VERSION AF202345.1 GI:11068678
KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 483)

AUTHORS Horn Davies, C.E., Davis, R.M. and Reidhaar-Olson, J.F.
TITLE Identification of toxicant-induced gene expression changes in HepG2
cells by differential display-PCR

JOURNAL Unpublished (2000)

COMMENT Contact: Horn Davies CE

Affymax Research Institute

3410 Central Expressway, Santa Clara, CA 95051, USA
Email: caroline.davies@affymax.com.

FEATURES Location/Qualifiers

SOURCE

1..483

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="128-7"

/clone_lib="Homo sapiens HepG2"

/cell_line="HepG2"

BASE COUNT 145 a 99 c 104 g 135 t
ORIGIN

alignment_scores:

Quality: 33.00 Length: 33
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-784-340-2 x AF202345 ..

Align seg 1/1 to: AF202345 from: 1 to: 483

443 LeuSerArgIleHisAspGlnProValLysProLeuAspArgAlaVala 459
|||||
104 TTATCAAGAAATTCATCATCAACGATCAACGCGCTGTGATGACGACGT 153
459 lPhetPrlleGluPheValMetArgHisLysGlyAlaLysHisLeuArg 475

|||||
154 CTCTGGATTGAGTTTGTGATGCGCCATAAGAGCCACCTTCGG 202
seq_name: gb_est83:BF080461

seq_documentation_block:

LOCUS BF080461 557 bp mRNA EST 18-OCT-2000
DEFINITION 231360 MARCH 2Pig Sus scrofa cDNA 5', mRNA sequence.
ACCESSION BF080461

VERSION BF080461.1 GI:10874291

KEYWORDS EST.

SOURCE pig.

ORGANISM Sus scrofa

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.

REFERENCE 1 (bases 1 to 557)

Fahrenkrug, S.C., Freking, B.A., Rohrer, G.A., Smith, T.P.L., Casas, E.,
Stone, R.T., Heston, M.P., Grosse, W.M., Bennett, G.A., Laegreid, W.W.
and Keele, J.W.

Design and use of two pooled tissue normalized cDNA libraries for
EST discovery in swine

Unpublished (2000)

Contact: Smith TPL

USDA, ARS, US Meat Animal Research Center

PO Box 166, Clay Center, NE 68933-0166, USA

Tel: 402 762 4366

Fax: 402 762 4390

Email: smith@email.marc.usda.gov

Single pass sequencing. Bases called and alt-trimmed with phred

v0.980904.e. Vector identified by cross-match with the -minscore 18

and -mismatch 12 options.

PCR Primers

FORWARD: AGGAACACGATGACCAT

BACKWARD: GTTCCGACGACGACG

Plate: 52 row: K column: 13

Seq primer: ATTAGTGACACTATAG.

FEATURES

source

1..557

/organism="Sus scrofa"

/db_xref="taxon:9823"

/clone_lib="MARCH 2Pig"

/clone_type="pooled"

/lab_host="DH10B"

/note="Vector: PCMV SPORT6; Site.1: XbaI; Site.2: XhoI;
Library made from pooled tissue from testis, ovary,"

endometrium, hypothalamus, pituitary, and placenta."

BASE COUNT 137 a 128 c 138 g 154 t
ORIGIN

alignment_scores:

Quality: 33.00 Length: 33
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-784-340-2 x BF080461 ..

Align seg 1/1 to: BF080461 from: 1 to: 557

443 LeuSerArgIleHisAspGlnProValLysProLeuAspArgAlaVala 459
|||||
251 TTATCAAGAAATTCACACGATCAACCTGTGAAGCCCTGGACGGCGGT 300
459 lPhetPrlleGluPheValMetArgHisLysGlyAlaLysHisLeuArg 475
|||||
301 CTCTGGATTGAGTTTGTGATGCGCCACAAAGAGCCACGACCTGCGG 349

seq_name: gb_est109:t67857

seq_documentation_block:

LOCUS t67857 403 bp mRNA EST 22-FEB-1995
DEFINITION yc27h03.r1 Stratiagene liver (#937224) Homo sapiens cDNA clone

IMAGE:81941 5' similar to gb:U05428 UDP-GLUCURONOSYLTRANSFERASE 2B7
 PRECURSOR, MICROSOMAL (HUMAN); mRNA sequence.

ACCESSION
 VERSION T67857
 KEYWORDS T67857.1 GI:679005

SOURCE
 ORGANISM human.
 Homo sapiens

REFERENCE
 AUTHORS 1 (bases 1 to 403)
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

Hillier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chapell, B.,
 Chissoe, S., Dietrich, N., Dubuque, T., Favello, A., Gish, W., Hawkins,
 M., Hultman, M., Kucaba, T., Lacy, M., Le, M., Le, N., Mardis, E., Moore,
 B., Morris, M., Parsons, J., Prange, C., Rifkin, L., Rohlfing, T.,
 Schellenberg, K., Soares, M.B., Tan, F., Thierry-Mieg, J., Trevaskis, E.,
 Underwood, K., Wohlmann, P., Waterston, R., Wilson, R. and Marra, M.
 Generation and analysis of 280,000 human expressed sequence tags
 Genome Res. 6 (9), 807-828 (1996)

TITLE
 JOURNAL MEDLINE
 COMMENT 97044478
 Contact: Wilson RK
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: estewatson.wustl.edu

Insert Size: 502

High quality sequence stops: 309 Source: IMAGE Consortium, LNL. This
 clone is available royalty-free through LNL; contact the IMAGE
 Consortium (info@image.llnl.gov) for further information.

Insert Length: 502 Std Error: 0.00
 Seq primer: M13RP1
 High quality sequence stop: 309.

FEATURES
 source Location/Qualifiers

1. 403
 /organism="Homo sapiens"
 /db_xref="GDB:498998"
 /db_xref="taxon:9606"
 /clone="IMAGE:81941"
 /clone.lib="Stratagene Liver (#937224)"
 /sex="male"
 /dev_stage="49 years old"
 /lab_host="SOLR cells (kanamycin resistant)"
 /note="Organ: liver; Vector: pBluescript SK; Site 1: EcoRI
 ; Site 2: XhoI; Cloned unidirectionally. Primer: Oligo
 dt. Hepatectomy from normal male caucasian. Average insert
 size: 1.1 kb; Uni-ZAP XR Vector: -5' adaptor sequence: 5'
 GAATTCGGCAGCAG 3' -3' adaptor sequence: 5'
 CTCGAGTTTTTTTTTTTTTTT 3'"
 BASE COUNT 114 a 72 c 94 g 123 t
 ORIGIN

alignment_scores:
 Quality: 28.00 Length: 28
 Ratio: 1.000 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
 US-09-784-340-2 x T67857 ..

Align seg 1/1 to: T67857 from: 1 to: 403

448 HtASpGlnProValLysProLeuAspArgAlaValAlaPheTrpIleGluPh 464
 |||||||
 58 CATGATCAACACAGTGAAGCCCTGGATCGAGCAGATCTTGATGATATT 107
 |||||||
 464 eValMetArgHisLysGlyValAlaLysHisLeuArg 475
 |||||||
 108 TGTCAATCGCCACAAAGAGACTAAACACCTTCGG 141

This Page Blank (uspto)